

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 1999, 23:01:47 ; Search time 360.53 Seconds
(without alignments)
1506.258 Million cell updates/sec

Title: US-09-205-015-3

Perfect score: 356

Sequence: 1 tcgacctctggaaacctatc.....atctggagctgaagaattc 356

Scoring table: IDENTITY_NUC

Searched: 2002476 seqs, 762712212 residues

Database : EST.*

- 1: em_est1.*
- 2: em_est2.*
- 3: em_est3.*
- 4: em_est4.*
- 5: em_est5.*
- 6: em_est6.*
- 7: em_est7.*
- 8: em_est8.*
- 9: em_est9.*
- 10: gb_est1.*
- 11: gb_est2.*
- 12: gb_est3.*
- 13: gb_est4.*
- 14: gb_est5.*
- 15: gb_est6.*
- 16: gb_est7.*
- 17: gb_est8.*
- 18: gb_est9.*
- 19: gb_est10.*
- 20: gb_est11.*
- 21: gb_est12.*
- 22: gb_est13.*
- 23: gb_est14.*
- 24: gb_est15.*
- 25: gb_est16.*
- 26: gb_est17.*
- 27: gb_est18.*
- 28: gb_est19.*
- 29: gb_est20.*
- 30: gb_est21.*
- 31: gb_est22.*
- 32: gb_est23.*
- 33: gb_est24.*
- 34: gb_est25.*
- 35: gb_est26.*
- 36: gb_est27.*
- 37: gb_est28.*
- 38: gb_est29.*
- 39: gb_est30.*
- 40: gb_est31.*
- 41: gb_est32.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB ID	Description
1	36.4	10.2	539 10	R18712 yf98f07.r1
2	35.2	9.9	895 26	AA788958 ag40d08.s

3	34.2	9.6	371 10	R26937
C 4	33.6	9.4	324 10	R51962
5	33.6	9.4	506 14	AA083476
C 6	33.6	9.4	361 18	AA373695
7	33.6	9.4	387 23	AI200590
8	33.6	9.4	401 23	AI247873
9	33.6	9.4	405 23	AI248529
10	33.6	9.4	376 25	AA676940
11	33.6	9.4	388 25	AA701392
12	33.6	9.4	413 29	AI049954
13	33.6	9.4	417 29	AI087130
14	33.6	9.4	353 29	RIC00677A
15	33.6	9.4	394 30	AI050745
16	33.6	9.4	388 30	AI080585
17	33.6	9.4	450 30	AI086664
18	33.6	9.4	436 30	AI086938
19	33.6	9.4	406 31	AI167997
20	33.4	9.4	582 12	N20987
21	33.4	9.4	357 30	AI056745
22	33.2	9.3	443 17	HSPD04245
23	32.8	9.2	237 13	W99245
24	32.8	9.2	554 15	AA081405
C 25	32.6	9.2	337 17	AA303662
C 26	32.2	9.0	467 10	T97491
27	32	9.0	385 12	N21587
28	32	9.0	586 15	AA155782
C 29	32	9.0	432 19	AA130208
30	32	9.0	521 19	AA133991
31	32	9.0	497 30	AI095618
32	31.8	8.9	606 12	N31902
33	31.8	8.9	423 14	W51911
34	31.8	8.9	425 20	AA462732
35	31.8	8.9	583 21	AA521429
36	31.8	8.9	500 31	AI184318
37	31.6	8.9	549 15	AA147860
38	31.6	8.9	427 24	AA423951
39	31.6	8.8	459 12	N24635
40	31.2	8.8	599 13	W22385
41	31.2	8.8	396 27	AA187138
C 42	31	8.7	477 13	N43182
43	31	8.7	427 22	AA565918
C 44	31	8.7	465 26	AA776725
C 45	30.8	8.7	556 31	C98110

ALIGNMENTS

RESULT 1

LOCUS	R18712	539 bp	mRNA	EST	14-APR-1995
DEFINITION	Yf98f07.r1 Homo sapiens CDNA clone 30551 5' similar to SP:MEC2_RAT Q00566 METHYL-CPG-BINDING PROTEIN 2 ;				
ACCESSION	R18712				
NID	q772322				
KEYWORDS	EST.				
SOURCE	human clone-30551 library-Soares infant brain lNIB vector-Lafmid BA host-DH10B (ampicillin resistant) primer-M13RPI Rsite1-Not I Rsite2-Hind III Whole brain from a 73 days post natal female. 1st strand CDNA was primed with a Not I - oligo(dT) primer [5', AACTGGAAGAAATTCGCGCGCAGCAATTTTTTTTTTTT 3']; double-stranded CDNA was ligated to Hind III adaptors (Pharmacia), digested with Not I and directionally cloned into the Not I and Hind III sites of the Lafmid BA vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo.				

ORGANISM	Homo sapiens
REFERENCE	Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS	1 (bases 1 to 539) Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,


```

1. 500
/organism="Homo sapiens"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dt.
Umbilical vein endothelial cells, passaged once. Average
insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor

```

96026280
MEDLINE
COMMENT
Other_ESTs: THC165970
Contact: Kerlavage, AR
BioInformatics
The Institute for Genomics

FEATURES

Location/Qualifiers
1 207


```
primer [5']
TGTTCACCAATCTGAAGTGGAGCGCGCATTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified p77T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo.
/db_xref="taxon:9606"
/clone="IMAGE:1681158"
/clone_lib="Soares_senescent_fibroblasts_NBHSF"
/tissue_type="senescent_fibroblast"
/lab_host="DH10B (ampicillin resistant)"

BASE COUNT 107 a 103 c 99 g 108 t
ORIGIN

Query Match 9.4%; Score 33.6; DB 29; Length 417;
Best Local Similarity 48.0%; Pred. No. 4.2;
Matches 96; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

Qy 95 aaacacttgaggagcagtaactgcccacacatgactcagtgcttcttgaggccaca 154
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 168 AAAAATAAAGGAAGCAGAAATCTGCTCAATGAGTAACACAGTCACCTACACTCCAAA 227
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 155 ggaacttcagtcactcctgtgggtggagggtgggacagggaagggtgaatggtact 214
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 228 TAACCCATGACACTGTTCTTGGGGAGGAGATGCCAGGCCAGGAGATGGGCCACAT 287
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 215 gctgattacaacctgtgtgctgcctccctcctctgttatctgagaggagccatgc 274
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 288 GATGGGGGACATTCACCTGCCACCTGCAGCACCCCTGTAAGTGGGGAGGGTGGCCAGTG 347
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 275 ccaaatgtttcacagccag 294
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 348 CCACAGTGGACTGTGAGATG 367
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
R1CC0677A 353 bp mRNA EST 15-JUL-1998
LOCUS
DEFINITION
R1CC0677A, partial sequence (C0677A), mRNA sequence.
ACCESSION
D15464
NID
g286657
KEYWORDS
EST; EST (expressed sequence tag).
SOURCE
Oryza sativa (strain:Nipponbare) callus cDNA to mRNA.
ORGANISM
Eukaryota; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta;
Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.
REFERENCE
Sasaki,T., Song,J., Koga-Ban,Y., Matsui,E., Fang,F., Higo,H.,
Nagasaki,H., Hori,M., Miya,M., Murayama-Kayano,E. et.al.
Toward cataloging all rice genes: large-scale sequencing of
randomly chosen rice cDNAs from a callus cDNA library
Plant J 6 (4), 615-624 (1994)
JOURNAL
95078950
MEDLINE
2 (bases 1 to 353)
AUTHORS
Sasaki,T. and Minobe,Y.
TITLE
Rice cDNA from callus
JOURNAL
Unpublished (1993)
REFERENCE
3 (bases 1 to 353)
AUTHORS
Sasaki,T.
TITLE
Direct Submission
JOURNAL
Submitted (14-APR-1993) to the DDBJ/EMBL/GenBank databases. Takuji
Sasaki, National Institute of Agrobiological Resources, Rice Genome
Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@abrr.affrc.go.jp. Tel:0298-38-7441,
Fax:0298-38-7468)
PROJECT = 'RGP'.
COMMENT
Location/Qualifiers
1..353
/organism="Oryza sativa"
source
```

```
/strain="Nipponbare"
/db_xref="taxon:4530"
/tissue_type="callus"

BASE COUNT 64 a 62 c 138 g 88 t 1 others
ORIGIN

Query Match 9.4%; Score 33.6; DB 29; Length 353;
Best Local Similarity 48.9%; Pred. No. 4.1;
Matches 90; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

Qy 26 ccacagtccagcagcagcacatctgcccaagcagaggtgagcagcatcagctgg 85
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 84 CCATGCCGCGAAGGTACTAGATATGTACGCCGAGGTGGCGGTAGTGGAGGGGAG 143
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 86 ggtctgtgaaacacttgaggagcagataactggggccacacatgactcagtcctcttg 145
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 144 GTGGTGGAGGGTCCGAGGATGGTTCTGGATGGGATCCGGCTCTGGCTCTGGTATGGTC 203
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 146 agccacacaggaactctgagtcactcctgtgggggtgaggtggagacaggaagggtg 205
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 204 AGCCGCGGAGATCTTCCGGTGGAGCATATGTAGTGGAGCGGGGAGGTCAAGGTGGTG 263
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 206 aatg 209
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 264 GAGG 267
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 15
A1050745 394 bp mRNA EST 24-SEP-1998
LOCUS
DEFINITION
ox94c10.s1 Soares_senescent_fibroblasts_NBHSF Homo sapiens cDNA
clone IMAGE:1663986 3', mRNA sequence.
ACCESSION
A1050745
NID
g3307550
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 394)
AUTHORS
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 992 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 386.
Location/Qualifiers
1..394
/organism="Homo sapiens"
/notice="Vector: p713D (Pharmacia) with a modified
polylinker V_Type: phagemid; Site_1: Not I; Site_2: Eco
RI; 1st strand cDNA was primed with a Not I - oligo(dT)
primer [5']
TGTTCACCAATCTGAAGTGGAGCGCGCATTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified p77T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo.
/db_xref="taxon:9606"
/clone="IMAGE:1663986"
/clone_lib="Soares_senescent_fibroblasts_NBHSF"
/tissue_type="senescent_fibroblast"
/lab_host="DH10B (ampicillin resistant)"
source
```


BASE COUNT	105 a	91 c	93 g	105 t
ORIGIN				

[illegible]

Search completed: June 23, 1999, 23:01:51
Job time: 1635 sec

